MicroGlam: Microscopic Skin Image Dataset with Cosmetics

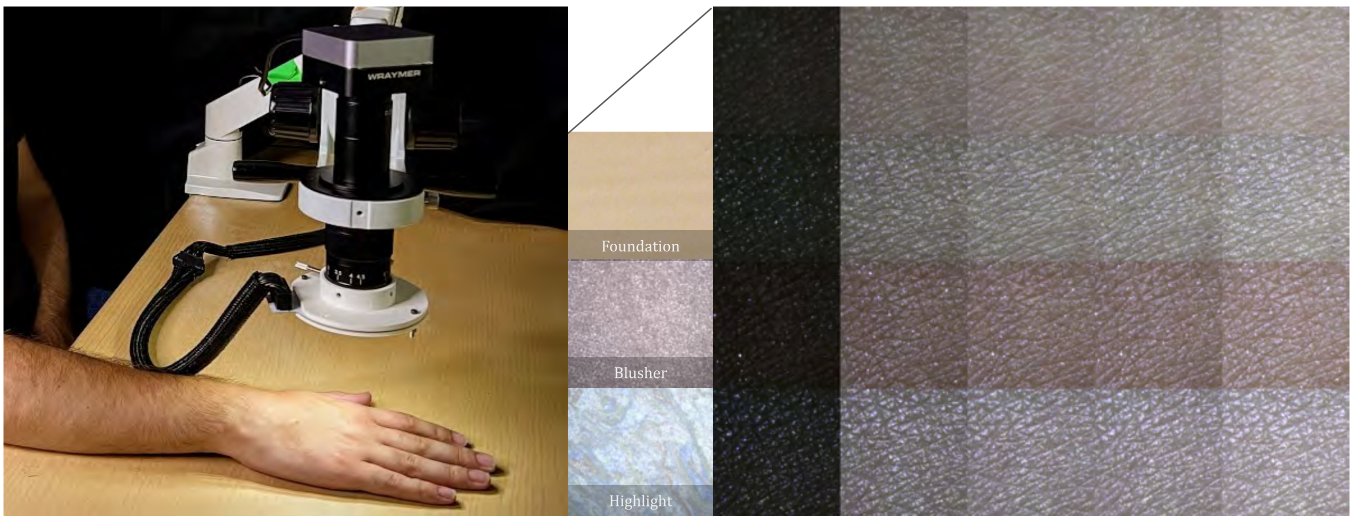
Toby Chong
tobyclh@gmail.com
The University of Tokyo
Japan

Alina Chadwick
alina.chadwick@gmail.com
Dartmouth College
USA

I-Chao Shen
ichaoshen@g.ecc.u-tokyo.ac.jp
The University of Tokyo
Japan

Haoran Xie
xie@jaist.ac.jp
JAIST
Japan

Takeo Igarashi
takeo@acm.org
The University of Tokyo
Japan



(a) (b) (c)

Figure 1: We developed (a) a capturing device consisting of a microscopic lens and 16 LEDs to capture a skin patch image dataset with cosmetics on the back of hand region. This dataset consists of $8mm^{*}8mm$ skin patch images under three cosmetic products (foundation, blusher, and highlighter). (c) We showed images captured under diverse lighting conditions for each cosmetic product.

ABSTRACT

In this paper, we present a cosmetic-specific skin image dataset. It consists of skin images from 45 patches (5 skin patches each from 9 participants) of size $8mm^{*}8mm$ under three cosmetic products (i.e., foundation, blusher, and highlighter). We designed a novel capturing device inspired by LightStage [Debevec et al. 2000]. Using the device, we captured over 600 images of each skin patch under diverse lighting conditions in 30 seconds. We repeated the process for the same skin patch under three cosmetic products. Finally, we demonstrate the viability of the dataset with an image-to-image translation-based pipeline for cosmetic rendering and compared our data-driven approach to an existing cosmetic rendering method [Kim and Ko 2018].

Authors' addresses: Toby Chong, tobyclh@gmail.com, The University of Tokyo, Japan; Alina Chadwick, alina.chadwick@gmail.com, Dartmouth College, USA; I-Chao Shen, ichaoshen@g.ecc.u-tokyo.ac.jp, The University of Tokyo, Japan; Haoran Xie, xie@jaist.ac.jp, JAIST, Japan; Takeo Igarashi, takeo@acm.org, The University of Tokyo, Japan.

2021. 0730-0301/2021/7-ART39 $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

CCS CONCEPTS

• **Computing methodologies → Computer graphics**; *Reflectance modeling*.

KEYWORDS

Skin capture, cosmetic rendering, skin dataset, makeup transfer

ACM Reference Format:
Toby Chong, Alina Chadwick, I-Chao Shen, Haoran Xie, and Takeo Igarashi. 2021. MicroGlam: Microscopic Skin Image Dataset with Cosmetics. *ACM Trans. Graph.* 38, 4, Article 39 (July 2021), 4 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

We examine existing cosmetic image datasets [Dantcheva et al. 2012; Gu et al. 2019; Li et al. 2018], which are often built to enable copying a general "look" from a face portrait with makeup to another without. They often consist of face portraits with and without makeup, mostly from different individuals. Utilizing the aforementioned datasets fundamentally entangles multiple challenging problems, such as separating the cosmetics from skin tone and lighting as well as separating the compounded effect of different cosmetic products when they are applied on top of each other or in multiple layers. The lack of specificity of the particular cosmetic product being applied onto portraits in most cosmetic datasets also limits the use of them to answer a very fundamental question, *"How would I look in this specific product?"*

We aim to simplify the problem to what we call **cosmetic rendering**, where a known cosmetic product is virtually applied onto an image of skin. We do so by creating the first cosmetic-specific skin image dataset ***MicroGlam***, a dataset of microscopic skin patches with and without cosmetics, designed and collected with the following goals in mind:

- contains the makeup and no-makeup data to accurately isolate the effect of a cosmetic of interest on skin,
- allows further processes to extract not only the diffuse color but also other visual parameters from the cosmetics,
- contains important metadata, including the cosmetic used, the method of application, and the skin tone of the subject.

We first developed a capturing and data processing framework, which contains 16 LEDs attached onto a microscope that lights the skin patch from 16 directions for accurate measurement. With this device, we captured skin patches at the back of the hand region, where cosmetic products (including foundation, blusher, and highlighter) are often tested in store. Using the proposed device, we captured data from 9 subjects aged 20-32. For each participant, we captured 600 images for each of five distinct skin patches at the back of the hand region. Each individual skin patch was captured under three different cosmetic products along with a reference capture in its natural, no-makeup state (Table 1). We demonstrated that our dataset can be leveraged to accurately render cosmetics onto an unseen skin patch with an unknown skin-tone and are of greater realism in comparison to the results generated by an existing cosmetic transfer work [Kim and Ko 2018].

2 RELATED WORK

2.1 Cosmetic datasets

Cosmetic datasets are often curated in the form of "unpaired" datasets [Gu et al. 2019; Li et al. 2018], where makeup and no-makeup portraits belong to different individuals. On the other hand, "paired" datasets, which include makeup and no-makeup portraits of the same person are rare. Dantcheva *et al.* [2012] created a paired dataset from YouTube makeup tutorials. The dataset is paired as it contains before and after makeup images of the same person, yet the faces are often under different lighting conditions and poses and the resolution of the videos cannot capture the subtle changes of individual products. Synthetic data is another popular approach to generate paired data by applying virtual makeup to no-makeup portraits [Sajid et al. 2018] and algorithmic makeup removal on makeup portraits [Yang et al. 2023]. Scherbaum *et al.* [2011] captured paired full face makeup data using a multi-light device similar to Lightstage, yet they only captured data from three people with no documentation regarding the cosmetic product(s) applied.

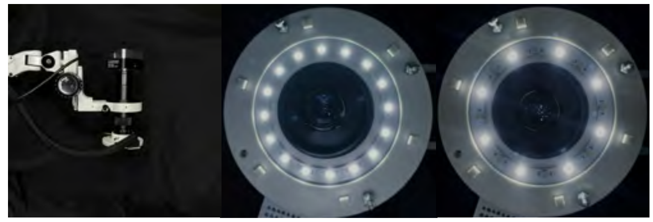


Figure 2: (a) Side view of our capturing device. (b)(c) Bottom view of the device with all and half of the LEDs selectively turned on, respectively. Each of the 16 LEDs surrounding the camera center can be switched on/off individually.

2.2 Detailed skin appearance capturing

Reproducing realistic skin appearance (both texture and geometry) is key to creating realistic virtual makeup. However, it remains challenging and requires specialized setups due to the complex physical and visual properties of skin, e.g.,non-rigid shape, microgeometry (wrinkles) and, subsurface scattering. One way to capture the microgeometry is to cast a mold onto the skin area [Haro et al. 2001]. However, this method cannot capture the interaction between the skin and cosmetics.

3 CAPTURING PROCEDURE AND DATA PROCESSING

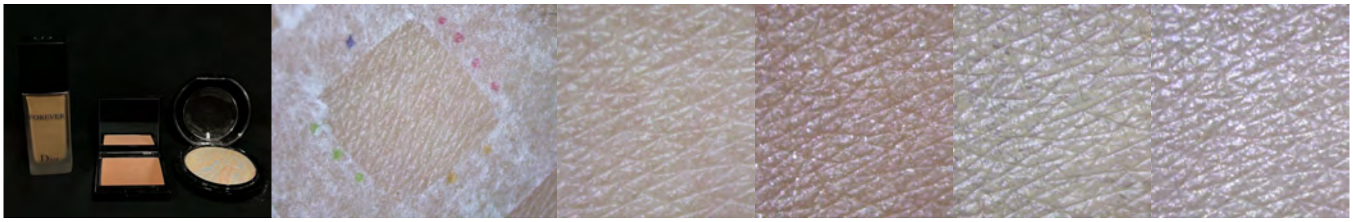
In this section, we discuss the design of our capturing device, the capturing procedure, and the data processing pipeline.

3.1 Capturing device

We had the first observation found that cosmetic products heavily interact with skin microgeometry (i.e., fine wrinkles and pores). We hence determined that microscopic data is required if we aimed to reproduce subtle changes to the skin. To this end, we based our capturing device on a digital microscope (Wraycam VEX120) with a native resolution of $1280^{*}960$. We capture skin at roughly $7,000$ pixel/cm (for scale, a 4K face portrait is roughly 200 pixel/cm). A 3D-printed ring light with 16 LEDs is placed uniformly surrounding the microscope center with a radius of $5mm$. The state of the lights (on/off) can be controlled individually to simulate complex lighting conditions from one or multiple directions. We controlled both the camera and the ring light at 20FPS to ensure the lights and camera at synchronized during our capturing procedure.

3.2 Subjects and cosmetic products

We recruited 9 subjects aged from 20-32. Their skin tones range from shade three to six in the MST scale [Monk 2023]. We sampled five patches from each of the subjects' hands. We captured their skin with highlighter (M.A.C. Mineralize Skinfinish Lightscapade), blusher (Laura Mercier blush color infusion 05 Shimmer), and foundation (Christian Dior Forever skin glow 2N) as well as a reference no-makeup capture.



(a) (b) (c) (d) (e) (e)

Figure 3: (a) Cosmetic products captured: foundation, blusher, and highlighter. (b) Raw image of a skin patch captured using our device of approximately $8mm^*8mm$. (c) Cropped and aligned version of (b). (d, e, f) The same skin patch but with blusher, foundation, and highlighter applied onto it respectively. We recommend enlarging the figures for a clearer viewing experience. Note how all makeup interacts with the skin microgeometry and affects more than just the diffuse color of the skin (e.g., foundation fills in the wrinkles of the skin and highlighter adds significant specular reflection onto the surface of skin).

(a) Subjects	9
(b) Cosmetic products	No-makeup (reference), foundation, blusher, highlighter
(c) Patches per person	five $8mm^*8mm$ skin patches
(d) Images per sequence	600 images for randomized lights and 16 images for one-light at a time
Total image count	$(a)^*(b)^*(c)^*(d)$ = 110880

Table 1: Dataset overview.

3.3 Capturing Procedure

At the beginning of the capture session, we prepared five registration markers per subject that consisted of a $18mm^*18mm$ patch of surgical tape with a $8mm^*8mm$ hole in each patch to expose the subject's skin. On the borders of the hole, we added coloured dots using marker pens for image alignment (Figure 3b). We applied the registration markers to one hand and captured a *sequence* of 616 images of the skin patch under each marker without cosmetic product, and repeat the process for the three cosmetic products. When applying highlighter or blush, we used designated brushes dipped in a small amount of product (enough to cover the patch) to apply the makeup. When applying foundation, we squeezed out product onto a palette and used a designated brush to apply small amounts of foundation to the patch of interest.

Within a sequence, we randomly varied both the intensity of the lights as well as the light combinations, with 3 to 8 of the 16 lights on at a time for 600 images plus additional 16 images with one light on at a time for a total of 616 images. Notice that the lighting conditions of all captured sequences are identical across patches and makeup, and only varies among different subjects. Each capture sequence took approximately 30 seconds, and the entire capturing session took around 45 minutes for five patches and four rounds of capturing.

3.4 Skin patch alignment

After capturing, we manually aligned all images of the same patch of skin with makeup to those without makeup. We manually identify correspondences in the first images of two sequences and compute the homography transformation. We applied the same transformation to all images from the makeup sequences to align to the no-makeup sequence.

4 APPLICATION - COSMETIC RENDERING

To demonstrate the potential of our dataset, we propose a cosmetic rendering application based on the image-to-image translation network [Zhu et al. 2017]. The goal is to estimate the appearance of a given skin patch when a specific cosmetic product is applied to it. Using the data we captured in Section 3, we constructed three sub-datasets for each target cosmetic products: *foundation*, *blusher*, and *highlighter*. Each sub-datasets consisted of paired skin patches captured both with and without makeup application. We merged captured skin patches from all subjects to improve the skin tone variation of each sub-datasets. For each target cosmetic product, we trained a separate network using the corresponding sub-dataset. During inference, given the appearance of the skin without any cosmetics, each trained model predicts the appearance of skin patches under the application of the corresponding cosmetic.

In addition, we compare our cosmetic rendering result with that generated using the method proposed by Kim *et al.* [2018] (*baseline*). The baseline method transfers cosmetic product applied onto a flat surface (Figure 5) to a skin patch. We chose the thickness parameter in the baseline method such that it minimizes the error between the baseline result and the ground truth.

In Figure 4, we demonstrate all five patches captured from a subject whose data is not used during training, under varied lighting conditions. We showed that using our dataset, our method can generate convincing cosmetic rendering results that are closer to the captured skin patches with cosmetics applied to them. In particular, our method preserves the microgeometry of the input skin patch, while generating the desired specular effects after applying different cosmetic products.

5 CONCLUSION

In this work, we presented MicroGlam, a dataset of microscopic skin patches with and without cosmetic product. This dataset is designed and collected to facilitate the development of cosmetic rendering. It contains a variety of patches under cosmetic products. We demonstrated the potential of using MicroGlam to perform

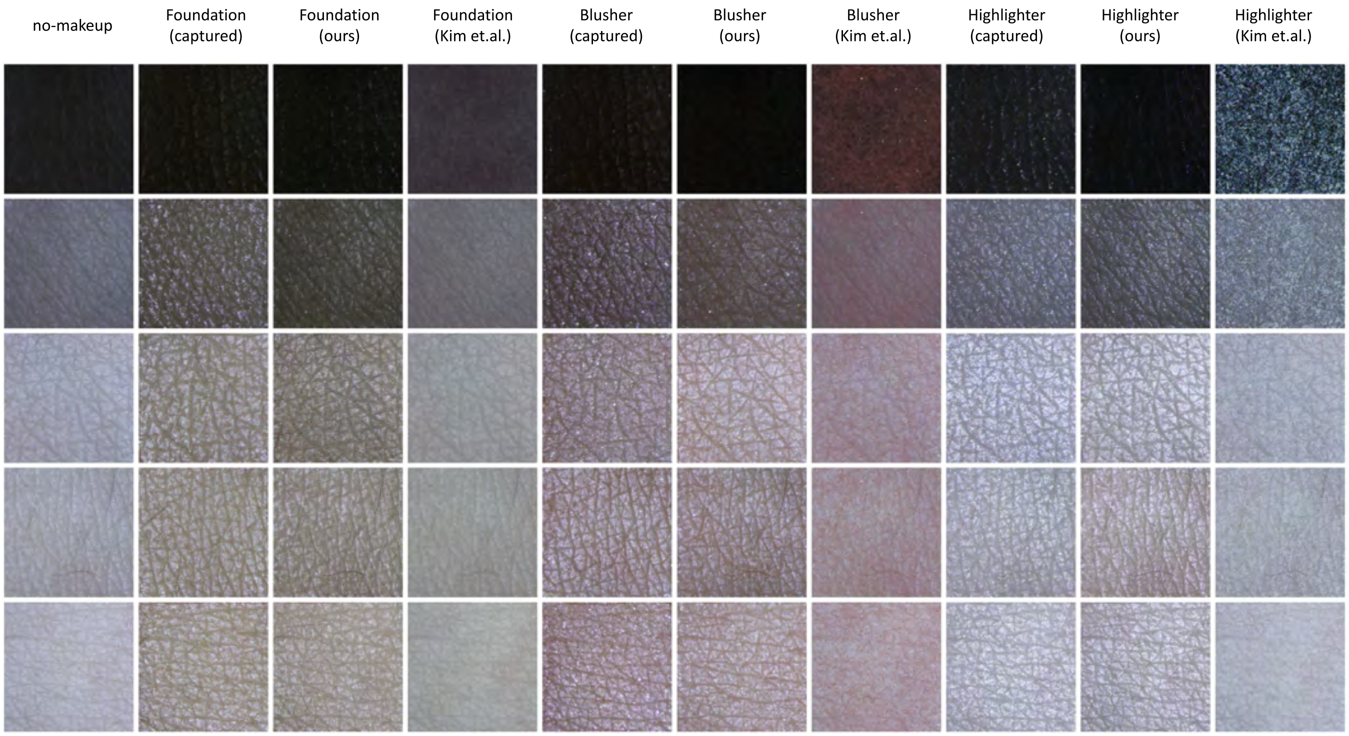


Figure 4: Cosmetic rendering results generated by our image translation-based approach (Section 4) and the baseline method ([Kim and Ko 2018]). We used the no-makeup images captured at different lighting conditions as input to generate all the results. For each cosmetic product, we provide a reference by presenting the captured skin patch with the exact same cosmetic product applied onto it.

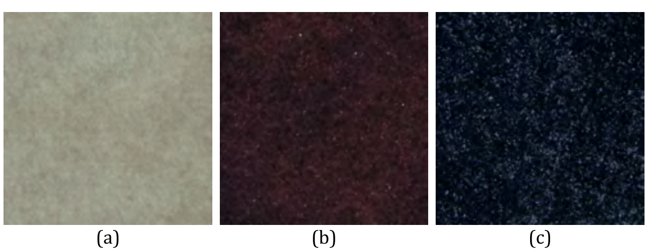


Figure 5: Following Kim *et al.*, we applied (a) foundation onto a white surface and (b) blusher and (c) highlighter onto a black surface. We use them as input to the baseline method. (zoomed-in images of different cosmetic products)

accurate cosmetic rendering. As future directions, we would expand on the diversity of our dataset in terms of cosmetics and skin-tones of subjects, and include skin data captured on the face.

6 ACKNOWLEDGMENT

We thank all subjects who participated in the capture sessions. This work is supported in part by Value Exchange Engineering, a joint research project between R4D, Mercari Inc. and the RIISE, JSPS Grant-in-Aid JP23K16921, JST AdCORP JPMJKB2302, and by NSF OISE 2153560.

REFERENCES

Antitza Dantcheva, Cunjian Chen, and Arun Ross. 2012. Can facial cosmetics affect the matching accuracy of face recognition systems?. In *BTAS*. IEEE, 391–398.

Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, and Mark Sagar. 2000. Acquiring the reflectance field of a human face. In *SIGGRAPH*. 145–156.

Qiao Gu, Guanzhi Wang, Mang Tik Chiu, Yu-Wing Tai, and Chi-Keung Tang. 2019. Ladn: Local adversarial disentangling network for facial makeup and de-makeup. In *ICCV*. 10481–10490.

Antonio Haro, Brian Guenter, and Irfan Essa. 2001. Real-time, photo-realistic, physically based rendering of fine scale human skin structure. In *Rendering Techniques 2001: Proceedings of the Eurographics Workshop in London, United Kingdom, June 25–27, 2001 12*. Springer, 53–62.

Goanghun Kim and Hyeong-Seok Ko. 2018. A Practical Approach to Physically-Based Reproduction of Diffusive Cosmetics. In *Computer Graphics Forum*, Vol. 37. 223–232.

Tingting Li, Ruihe Qian, Chao Dong, Si Liu, Qiong Yan, Wenwu Zhu, and Liang Lin. 2018. Beautygan: Instance-level facial makeup transfer with deep generative adversarial network. In *ACMMM*. 645–653.

Ellis Monk. 2023. The Monk Skin Tone Scale. (2023).

Muhammad Sajid, Nouman Ali, Saadat Hanif Dar, Naeem Iqbal Ratyal, Asif Raza Butt, Bushra Zafar, Tamoor Shafique, Mirza Jabbar Aziz Baig, Imran Riaz, and Shahbaz Baig. 2018. Data augmentation-assisted makeup-invariant face recognition. *Mathematical Problems in Engineering* 2018 (2018), 1–10.

Kristina Scherbaum, Tobias Ritschel, Matthias Hullin, Thorsten Thormählen, Volker Blanz, and Hans-Peter Seidel. 2011. Computer-suggested facial makeup. In *Computer Graphics Forum*, Vol. 30. 485–492.

Xingchao Yang, Takafumi Taketomi, and Yoshihiro Kanamori. 2023. Makeup Extraction of 3D Representation via Illumination-Aware Image Decomposition. In *Computer Graphics Forum*, Vol. 42. 293–307.

Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *ICCV*. 2223–2232.